Supplementary Materials

S1: Supplemental Methods

Sampling Process: simulating fishery-dependent data collection

For the *Preferential* sampling scenario, the fishing suitability raster resulted from a positive logistic function of the habitat suitability of the target species in the previous year (y-1). We used the y-1 suitability as fishermen often base their choice of where to fish on where they found the fish in the previous year (Sampson 1991). Therefore, we are using the habitat suitability in the previous year as a proxy of where the fishermen were likely to find the fish in the previous year for this analysis. We built the fishing suitability raster using beta βs (inflection points) for the logistic function of 0.7.

For the *Port* scenarios, fishing suitability was a function of a positive logistic response to target species habitat suitability in the previous year, with a beta=0.7, and a negative logistic response to distance from port. We built fishing suitability rasters for 8 different distance from port scenarios. Two scenarios simulated situations where fishermen were limited to fishing around northern ports in Washington (Westport, WA -124.114934, 46.911534) and Oregon (Garibaldi, OR, -124.292000, 43.383975), two scenarios simulated fishermen being limited to fishing around ports within the middle of the ROMS domain (Santa Cruz, CA (-122.001620, 36.965719), and Bodega Bay, CA (-123.050618, 38.334302)), two scenarios simulated fishing limited to fishing around a port in southern CA (San Diego Bay, CA (-117.1441, 32.6717)), and two scenarios included all ports. One of the scenarios for each port simulated an offshore fishery where fishing suitability was high up until about 300 miles from a port, and one scenario simulated a nearshore fishery where fishing suitability declines after about 50 miles from a port. The distance from each port to every cell in the raster was calculated using *distanceFromPoints* function in R. For the scenarios where more than one port was used, the lesser distance of each cell to the ports was used.

For the *Bycatch* scenario, we simulated the distribution of a turtle-like bycatch species which we used to impact fishing location suitability. The simulated bycatch species preferred warmer temperatures than our target species, exhibiting a unimodal response to SST (mean=25,

sd=10), and similar to the simulated prey species of our target species exhibited a positive logistic response to zooplankton abundance. The fishing location suitability was determined by a positive logistic function of target species habitat suitability in the previous year (with beta=0.7), and a negative logistic function of the bycatch species habitat suitability in the current year. This simulates a situation where fishermen actively avoid fishing in areas of high bycatch risk (high habitat suitability for bycatch species) but are still also attempting to fish in areas that have good habitat suitability for their target species (Smith et al. 2021, Hazen et al. 2018, Howell et al. 2008).

For the *Closed Area* scenario, fishing suitability was determined by multiplying the fishing suitability raster developed for the *Preferential* scenario with a beta=0.7 by a raster of 0s and 1s, where 0 values were for cells within the closed area. This effectively made it so cells within the closed area had a 0% chance of being sampled. We simulated fishing locations based on three different sizes of the closed area. The smallest closed area was a box with its corners at longitude and latitudes of -126, -118, 36, 41. The medium sized closed area had corners at -127, -118, 35, 42, and the largest closed area at -129, -118, 34, 43. The largest closed area was similar in size and location to the seasonally closed Pacific Leatherback Conservation Area (PLCA) off the US west coast; the PLCA is intended to reduce incidental bycatch of endangered leatherback turtles (*Dermochelys coriacea*) (Urbisci et al 2016).

Assessment of Climatic Bias in the Sampling Scenarios

We used two different metrics to obtain climatic bias and novelty, Cohen's d and Hellinger Distance (HD). Cohen's d is a measure of the difference between the means of two groups, and we calculated this using the below formula:

$$\mathrm{d} = \frac{(M1 - M2)}{Pooled\ SD}$$

Where M1 is the mean of the first sampling regime (e.g. SST in the partially sampled historical climate), and M2 is the mean for the second sampling regime being compared (e.g. SST in the full historical or future climate used for prediction), and the denominator is the pooled standard deviation. A value of d = 0.2 is generally considered a small difference or effect, 0.5 a medium, and 0.8 a large difference (Cohen 1988). The Hellinger distance is a measure of the difference

between two probability distributions (see Legendre & Legendre 2012) and we calculate this using the below formula for each environmental parameter:

$$H(P, Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^{k} (\sqrt{p_i} - \sqrt{q_i})^2}$$

Where p_i is the probability distribution of the environmental parameter of interest in the entire sampling domain and q_i is the probability distribution for that same environmental parameter for a particular sampling scenario. The HD measures how much information is contained in one distribution relative to another. With this metric the two distributions being compared become more similar as the difference in the proportion of sites at each value of the environmental covariate declines. HD values >0.5 have been proposed as a threshold of novelty, where the distributions become more dissimilar than they are similar (Johnson and Watson et al *in prep*).

Table S1: Variable used to simulate species spatially explicit distribution and fishery-dependent sampling scenario suitability rasters. μ is the mean and σ the standard deviation for the normal response curves. For the logistic response curves, α is the scale parameter which controls the slope of the curve (the growth rate), and β is the location parameter specifying the time when curve reaches the midpoint of the growth/decline trajectory.

Suitability Raster	Variable	Description	Parameter 1	Parameter 2	Distribution
HMS species archetype	SST	Sea surface temperature	μ = 17	σ = 4	normal
	MLD	Mid-layer depth	μ = 50	σ = 30	normal
	Prey presence	Preference for prey	α = -0.15	β = 0.4	logistic
	Biomass (kg)	Biomass if the species is present in a grid cell	log μ = 3.29	log (σ) = 0.26	Log normal
	Occurrence (0 or 1)	Occurrence as a function of habitat suitability	α = -0.07	β = 0.4	logistic
Prey species	SST	Sea surface temperature	μ = 14	σ = 7	normal
	zoo_200m	Zooplankton integrated over top 50m	α = -10	β = 45	logistic
Bycatch species	SST	Sea surface temperature	μ = 25	σ = 10	normal
	zoo_200m	Zooplankton integrated over top 50m	α = -6	β = 50	logistic
Preferential	HMS species presence t-1	HMS habitat suitability in the previous year	α = -0.05	β = 0.7	logistic
Distance - Nearshore scenarios	HMS species presence t-1	HMS habitat suitability in the previous year	α = -0.05	β = 0.7	logistic
	max distance	Controls how suitability declines in relation to distance from port	α = 50	β = 110	logistic
Distance - Offshore scenarios	HMS species presence t-1	HMS habitat suitability in the previous year	α = -0.05	β = 0.7	logistic
	max distance	Controls how suitability declines in relation to distance from port	α = 50	β = 480	logistic
Bycatch avoidance	HMS species presence t-1	HMS habitat suitability in the previous year	α = -0.05	β = 0.7	logistic
	Bycatch species presence	Habitat suitability of a simulated bycatch species	α = 0.05	β = 0.5	logistic

Table S2: Model formulation for the GAMs and BRTs fitted in the simulation

Model	Description	Model Component	R syntax
GAM_E	environmental covariates only	Pr(presence)	s(SST)+ s(MLD) + s(chl); family = binomial
		log(biomass)	s(SST)+ s(MLD) + s(chl); family = gaussian
GAM_E+ST	environmental covariates and spatiotemporal covariates	Pr(presence)	(SST) + s(MLD) + s(chl) + te(lat, lon, year, d=c(2,1), bs=c('ds','tp'), m=(c(1,.5),1), k=c(30,10))
		log(biomass)	(SST) + s(MLD) + s(chl) + te(lat, lon, year, d=c(2,1), bs=c('ds','tp'), m=(c(1,.5),1), k=c(30,10))
BRT_E	environmental covariates only	Pr(presence)	gbm.x = c(SST, MLD, chl-surface) gbm.y = 'pres', family = "bernoulli", tree.complexity = 3, learning.rate = 0.01, bag.fraction = 0.6
		log(biomass)	gbm.x = c(SST, MLD, chl-surface) gbm.y = 'log(biomass)', family = "gaussian", tree.complexity = 3, learning.rate = 0.01, bag.fraction = 0.6,
BRT_E+ST	environmental covariates and spatiotemporal covariates	Pr(presence)	gbm.x = c(SST, MLD, chl-surface, lat, long, year) gbm.y = 'pres', family = "bernoulli", tree.complexity = 3, learning.rate = 0.01, bag.fraction = 0.6
		log(biomass)	gbm.x = c(SST, MLD, chl-surface, lat, lon, year) gbm.y = 'log(biomass)', family = "gaussian", tree.complexity = 3, learning.rate = 0.01, bag.fraction = 0.6

Table S3: Deviance explained for the presence and abundance components of the GAMs and BRTs models for both those with only environmental covariates and the ones which included both environmental and space-time covariates.

Scenario	Presence Component				Abundance Component			
	Environment Only		Env + Spacetime		Environment Only		Env + Spacetime	
	GAM	BRT	GAM	BRT	GAM	BRT	GAM	BRT
Bycatch	45	49.8	51.7	51.4	35	41	38.7	44.1
Closed Area Large	53.1	57.4	60	62.9	37.7	40.8	37.7	42.3
Closed Area Medium	49.9	54.8	53.1	57	36.6	41.4	38.9	45
Closed Area Small	48.9	54.1	50.1	58.7	30.5	34.6	33.5	35.2
Port All Nearshore	27.3	36	27.4	31.7	29.6	33.9	30.3	34
Port All Offshore	41.4	47.3	46	48	28.1	33.7	28.7	35.9
Port Middle Nearshore	27.1	34.9	27.2	32.7	28.1	34.2	28.5	34.9
Port Middle Offshore	48.2	58.1	49.1	56.7	30.7	34.9	30.7	35.5
Port Northern Nearshore	25.6	27.96	25.9	28.2	36.2	40.6	36.3	42.1
Port Northern Offshore	24.6	28.5	25.5	28.7	29.4	35.7	29.4	36.6
Port Southern Nearshore	32.1	35.6	35.1	36	34.4	38.1	34.4	37.8
Port Southern Offshore	35.4	40.6	38.2	45.2	26.2	31.3	26.5	32.7
Preferential	47.9	52.9	52.4	56.2	32	36.1	37.8	37.4
Random	44.9	49.3	47.6	50.1	57.7	63.5	57.8	64.9

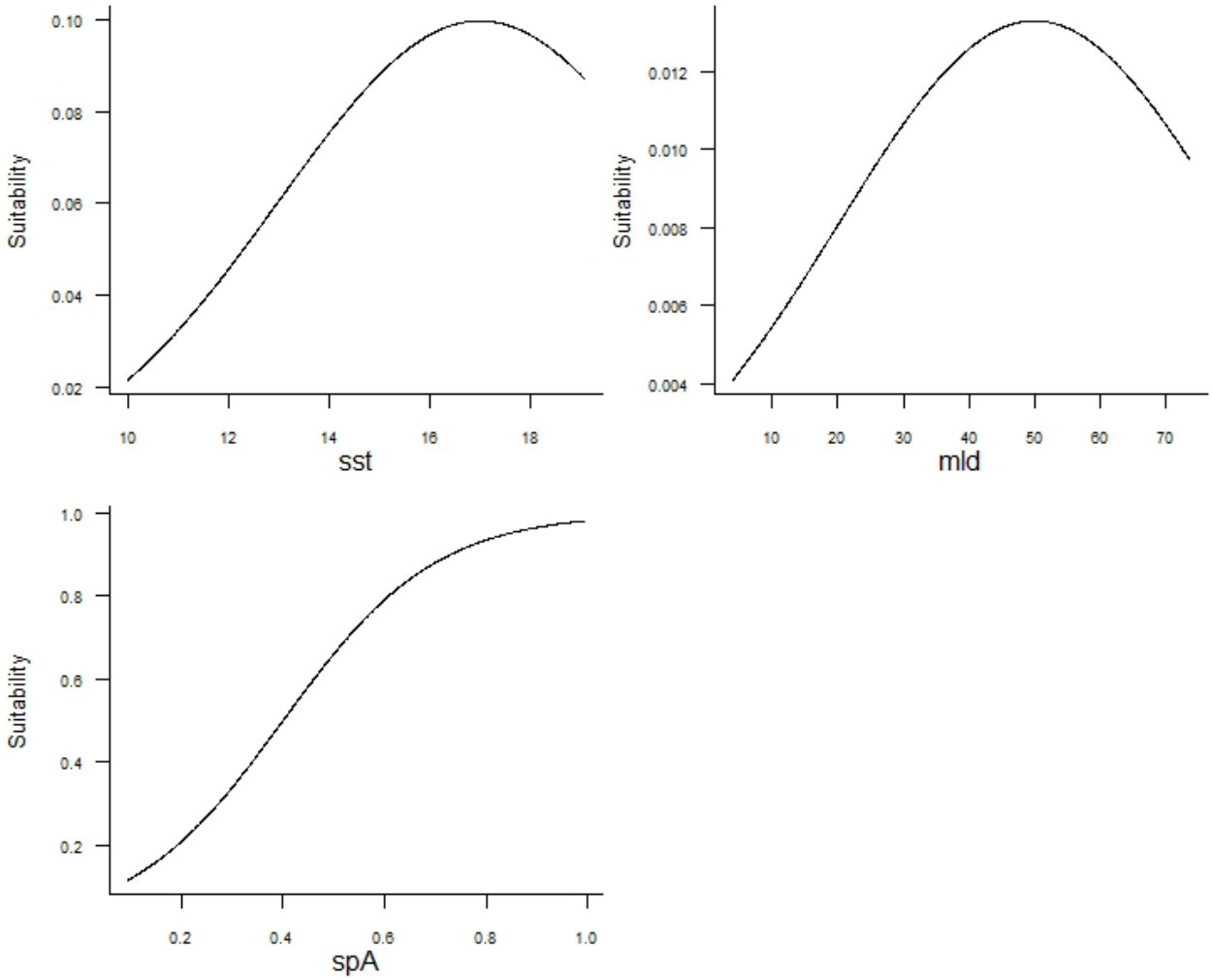

Figure S1. Virtual species response curves. ‘spA’ is the prey species, and represents the target species preference for its prey species.

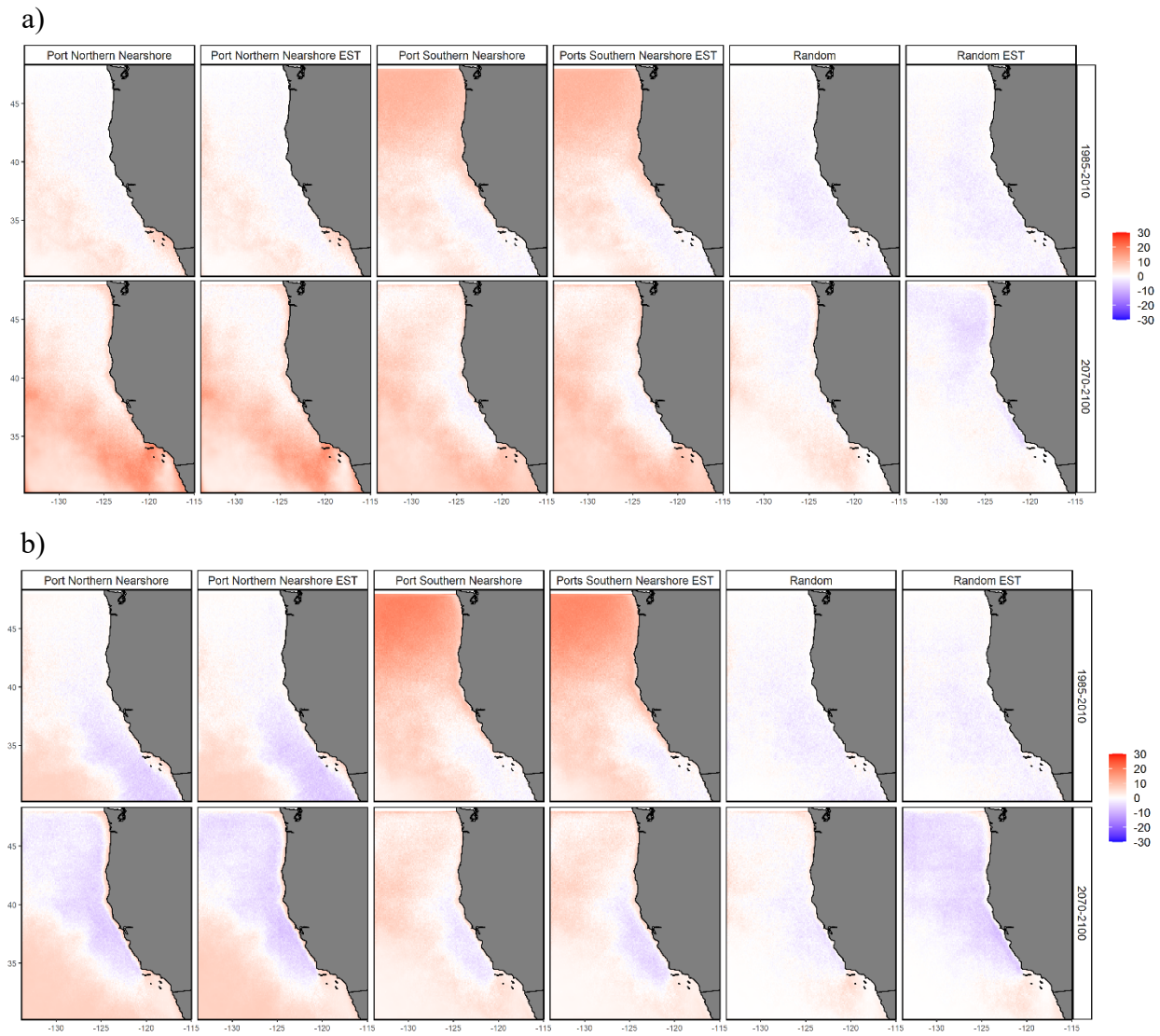

Figure S2: (a) Spatial residual (predicted – observed) maps for the GAM models, and (b) spatial residual maps for the BRT models for selected sampling scenarios fit Environmental covariates only, and Environmental and Space-Time covariates (EST). Only the historic training period (1985-2010) and late-century future period (2070-2100) are shown for simplicity.

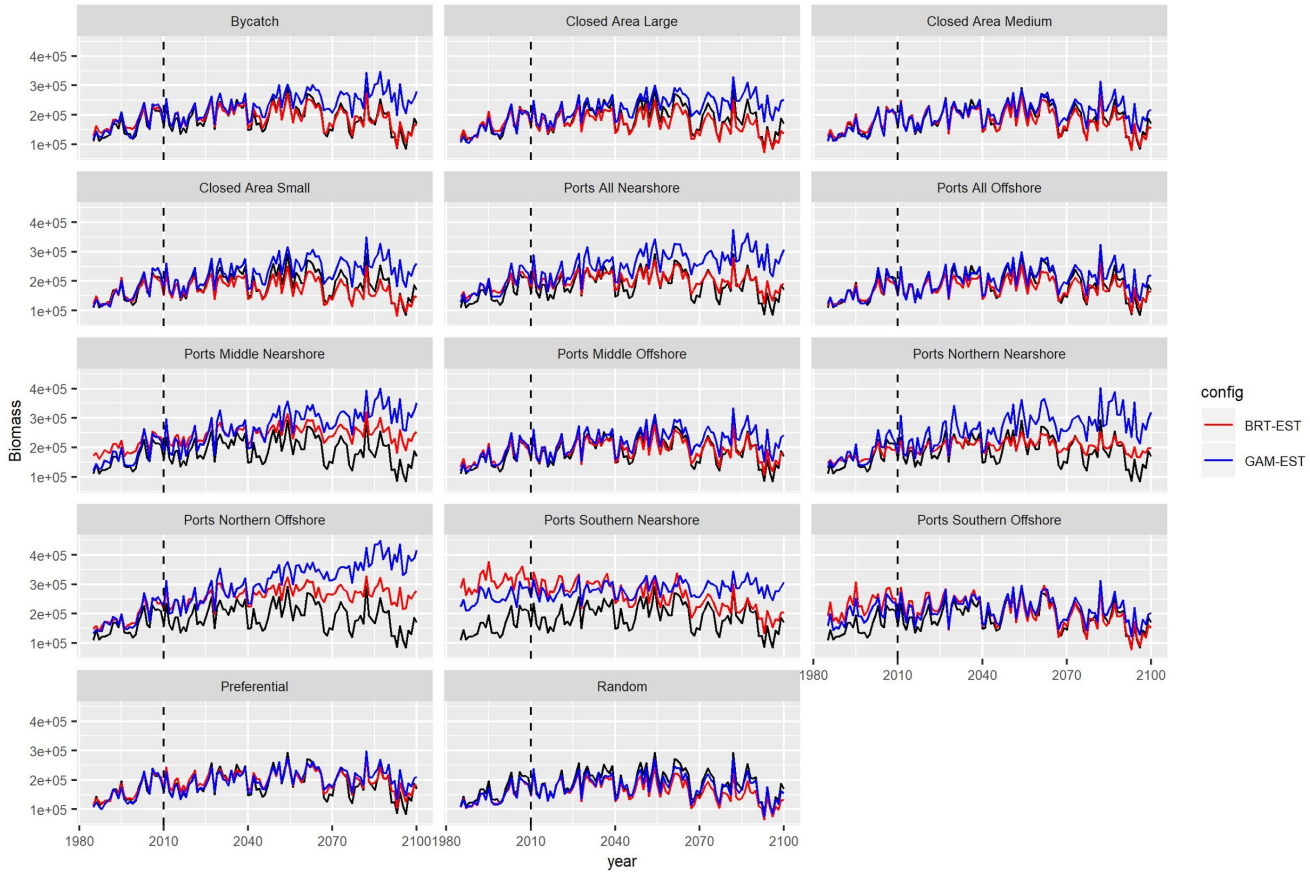

Figure S3: Biomass time series for models fit with space and time covariates as well as all three environmental covariates (MLD, SST, CHL).

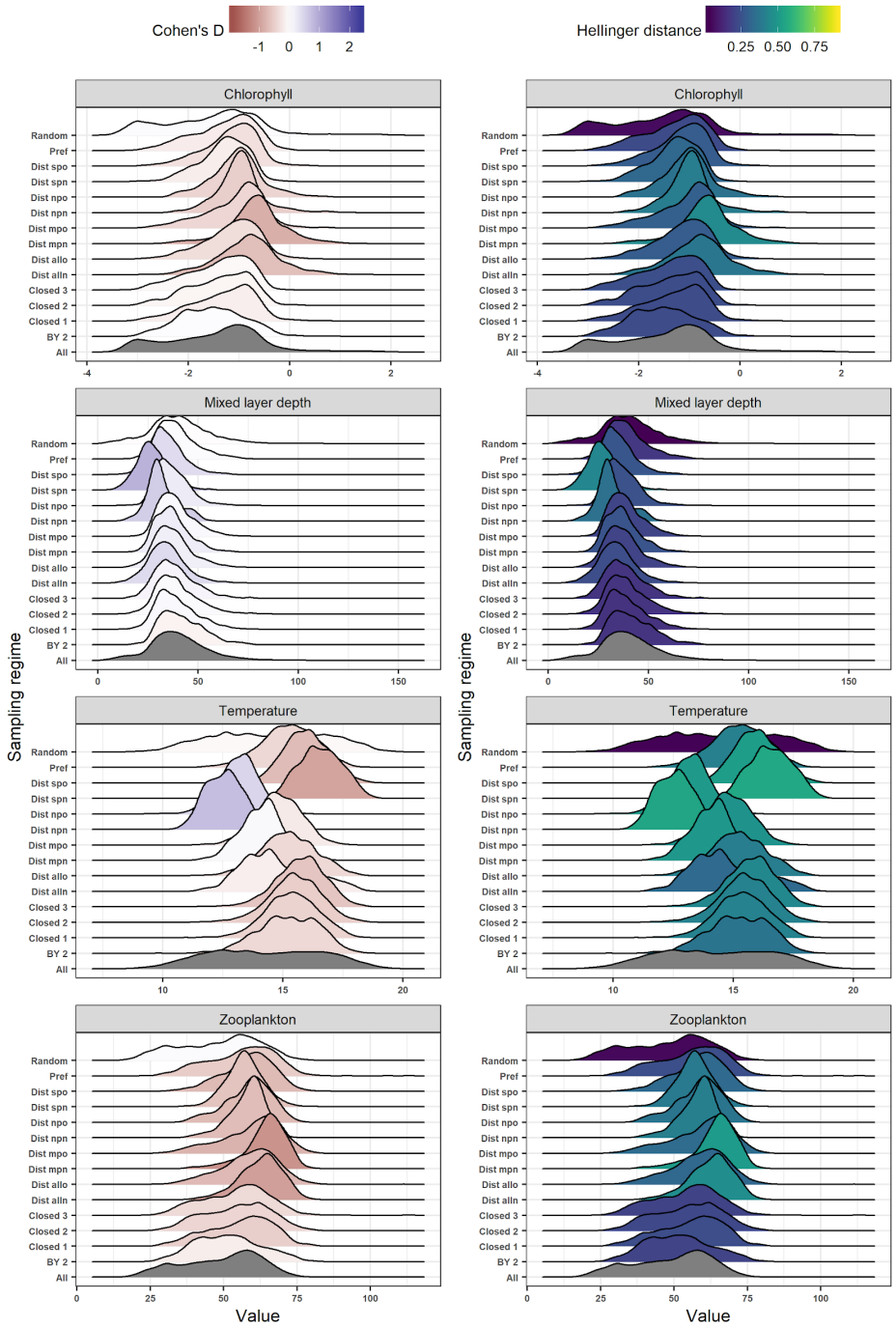

Figure S4: Distribution and climate bias of environmental variables in each sampling scenario for historical period (1985-2010). Colors indicate the magnitude of the climate bias of each sampling scenario, measured with Cohen's D (left panels) and Hellinger distance (right panels).

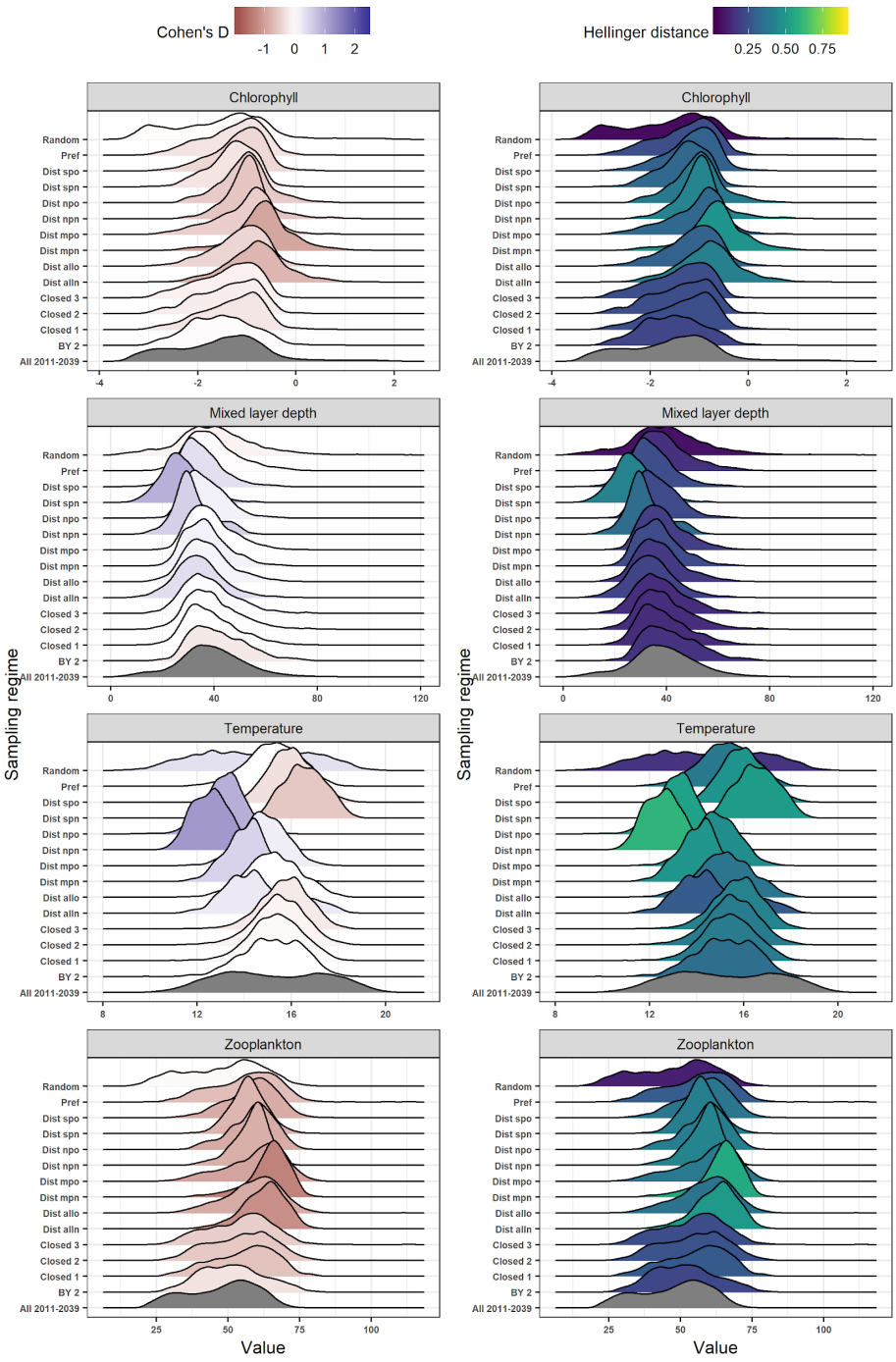

Figure S5: Distribution and climate bias of environmental variables in each sampling scenario for early-century period (2011-2039). Colors indicate the magnitude of the climate bias of each sampling scenario, measured with Cohen's D (left panels) and Hellinger distance (right panels).

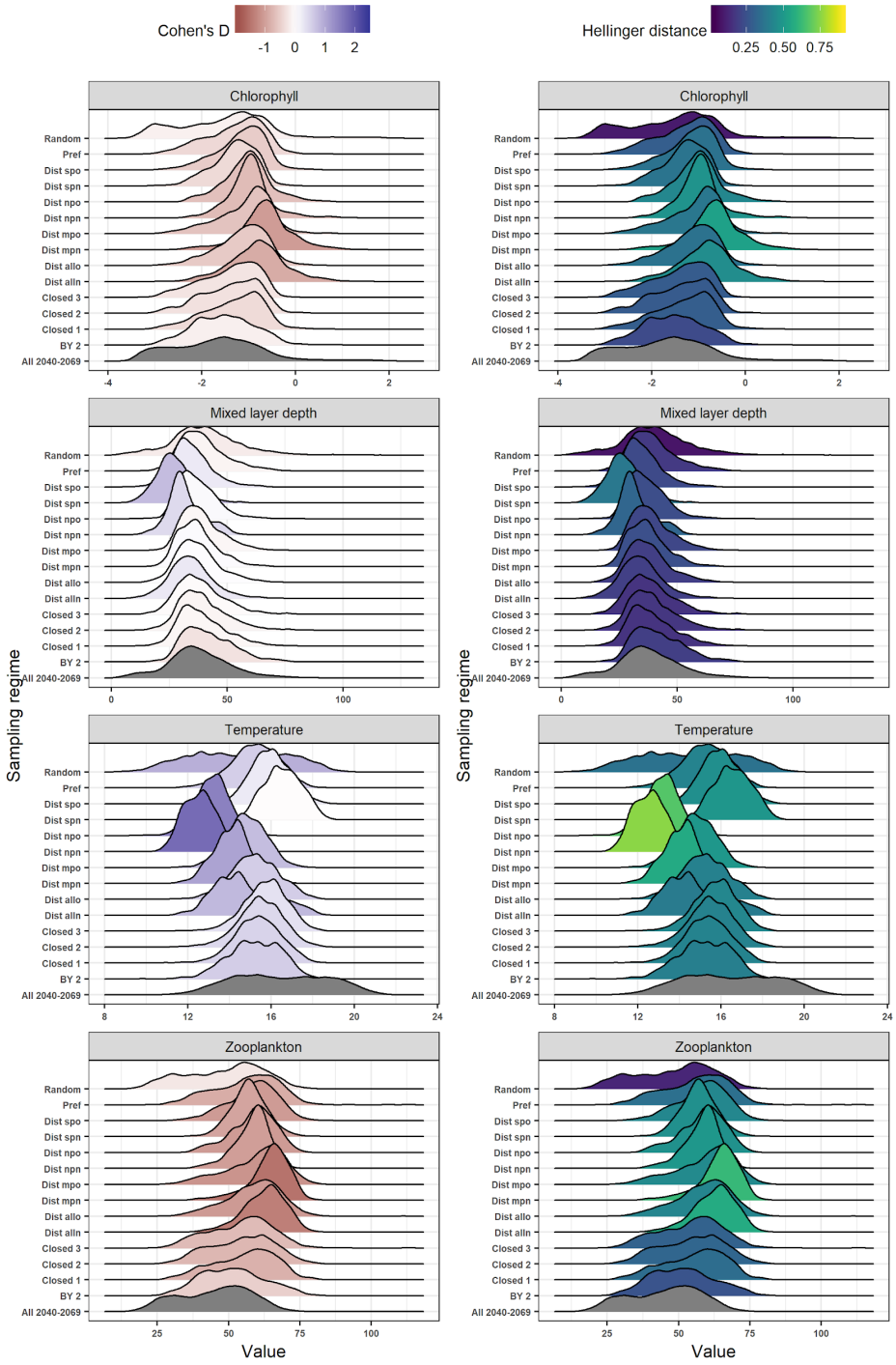

Fig S6: Distribution and climate bias of environmental variables in each sampling scenario for middle-century period (2040-2069). Colors indicate the magnitude of the climate bias of each sampling scenario, measured with Cohen's D (left panels) and Hellinger distance (right panels).

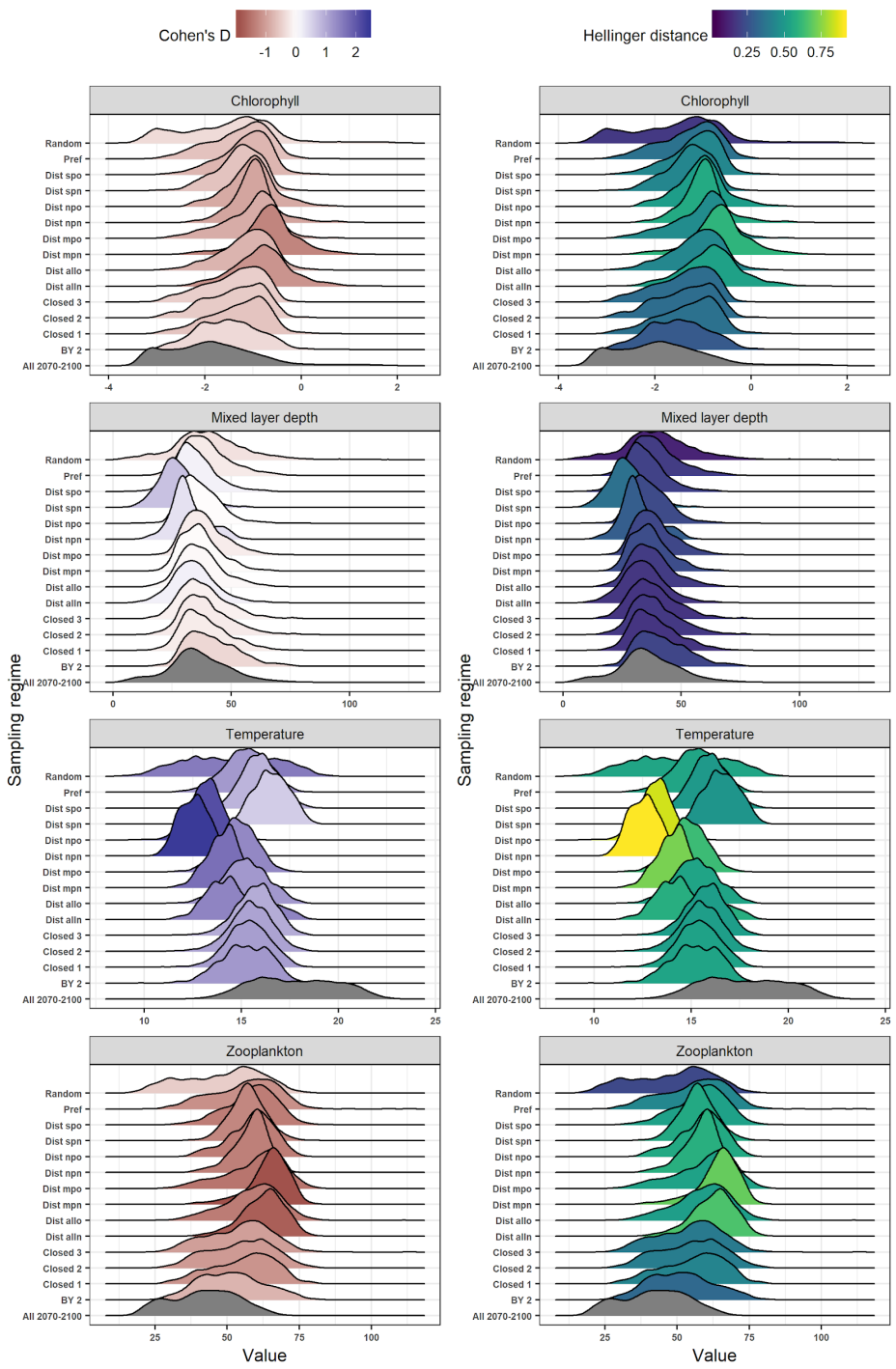

Fig S7: Distribution and climate bias of environmental variables in each sampling scenario for late-century period (2070-2100). Colors indicate the magnitude of the climate bias of each sampling scenario, measured with Cohen's D (left panels) and Hellinger distance (right panels).

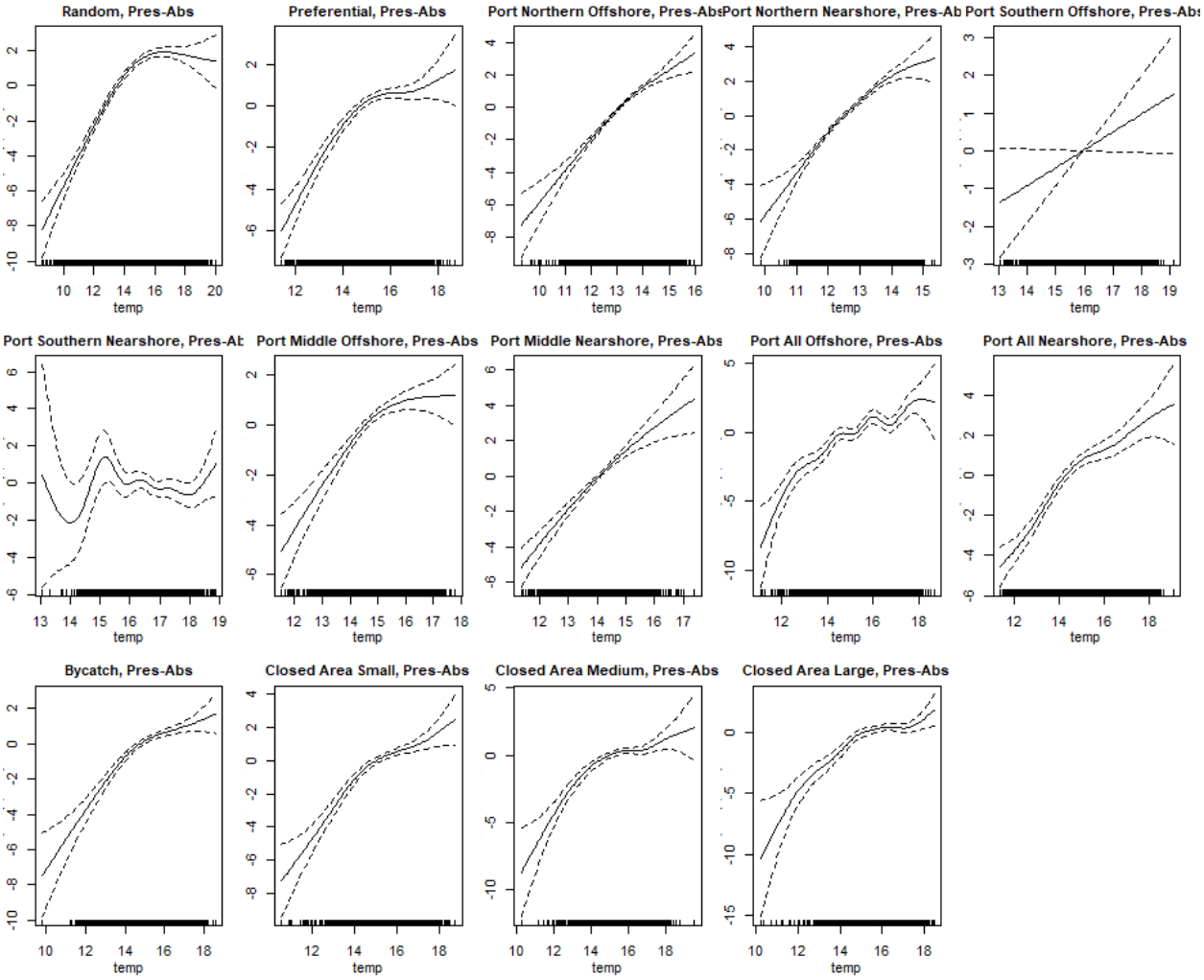

Figure S8. Sea surface temperature response curves for the binomial (occurrence, presence-absence) part of delta model (GAMs).

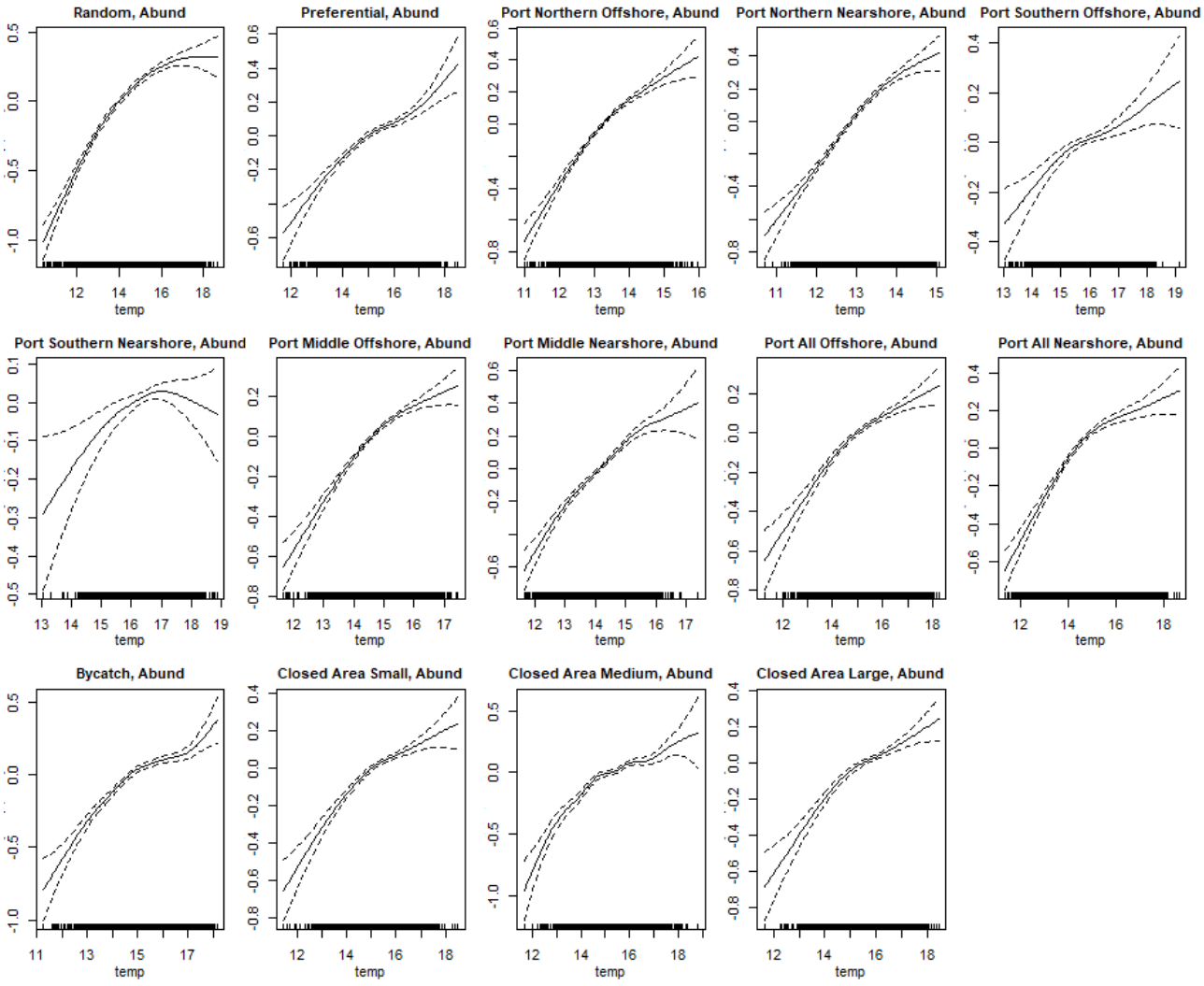

Figure S9. Sea surface temperature response curves for abundance part of delta model (GAMs).

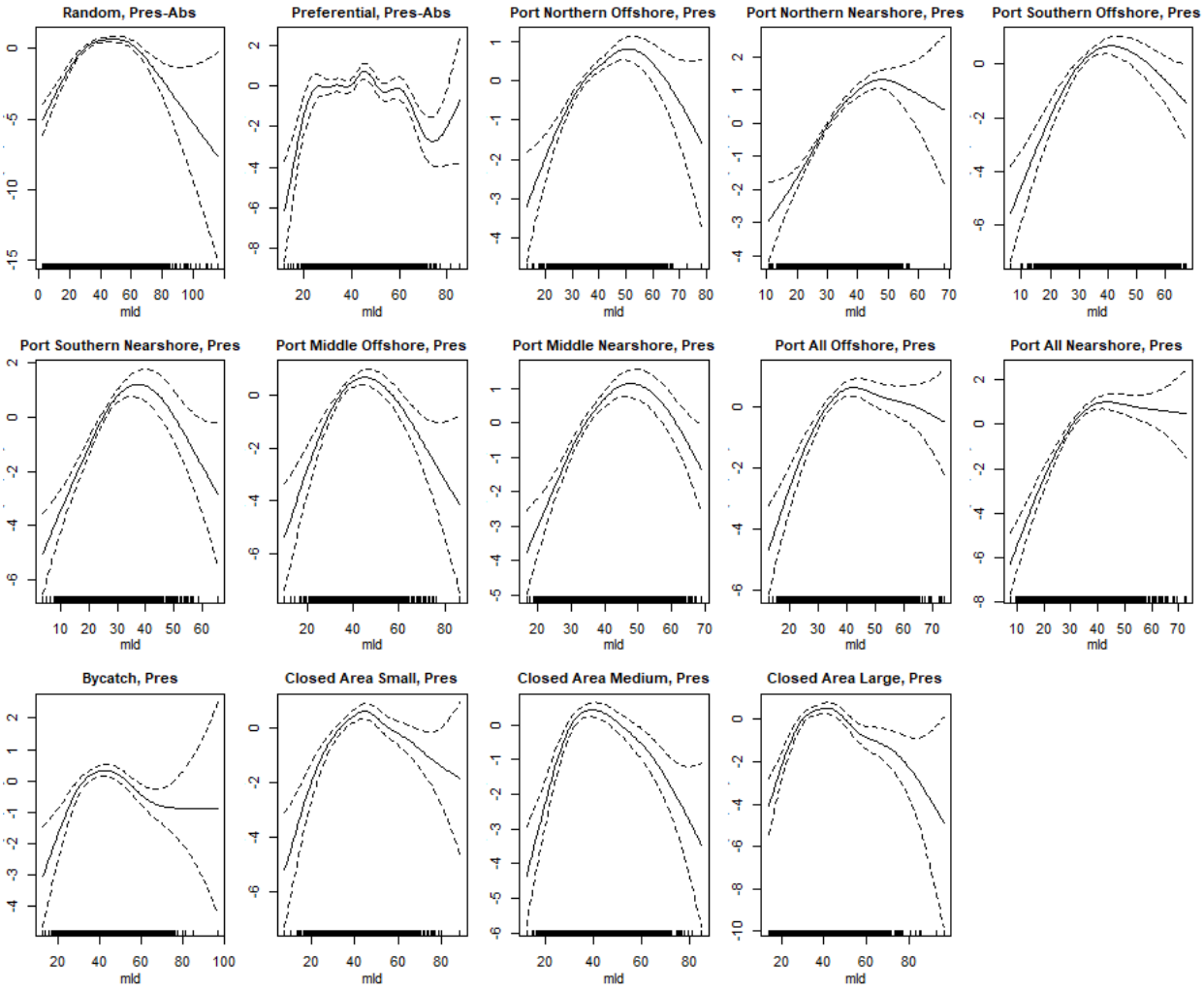

Figure S10. Mixed layer depth (mld) response curves for binomial (occurrence, presence-absence) part of delta model (GAMs).

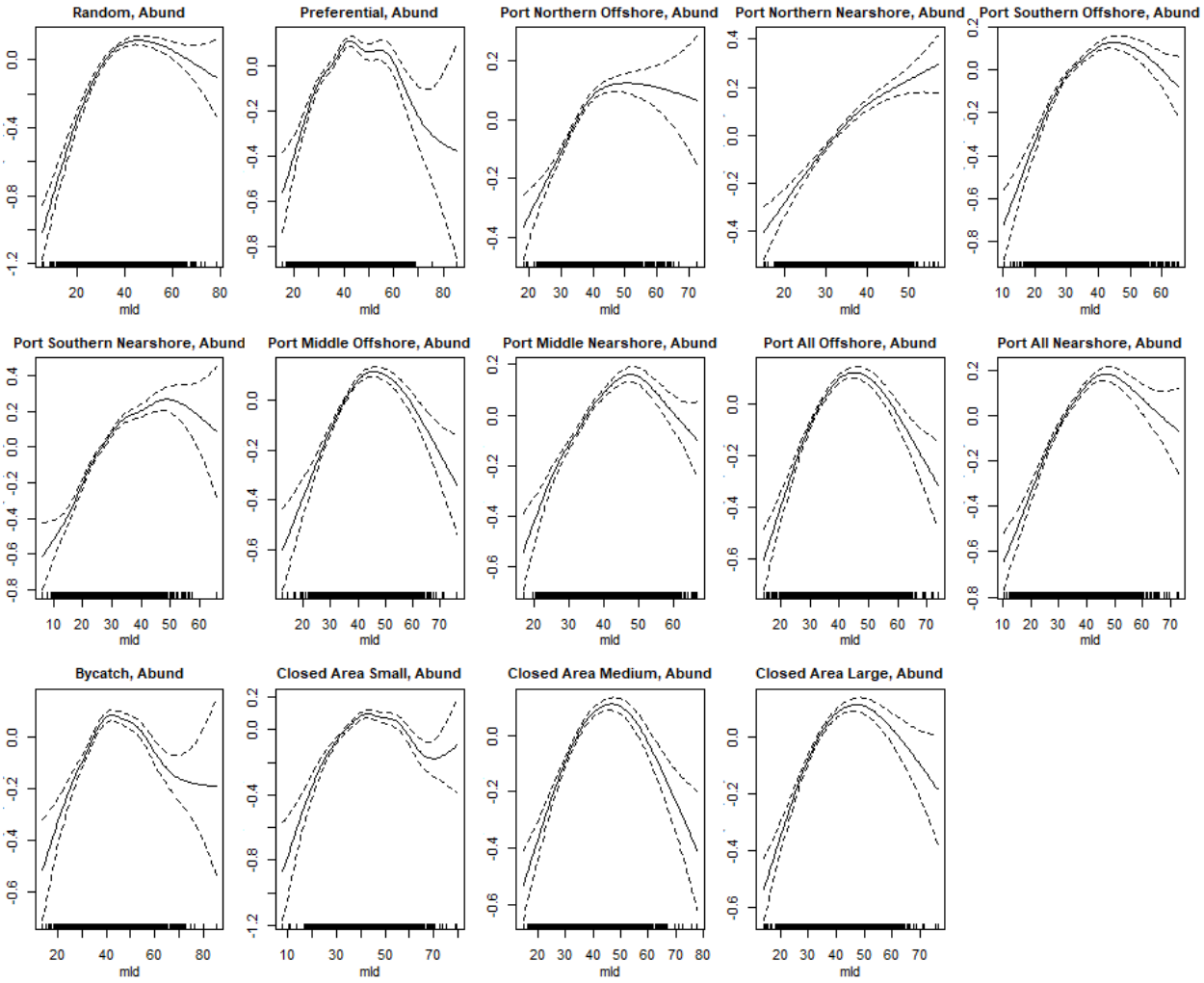

Figure S11. Mixed layer depth (mld) response curves for abundance part of delta model (GAMs).

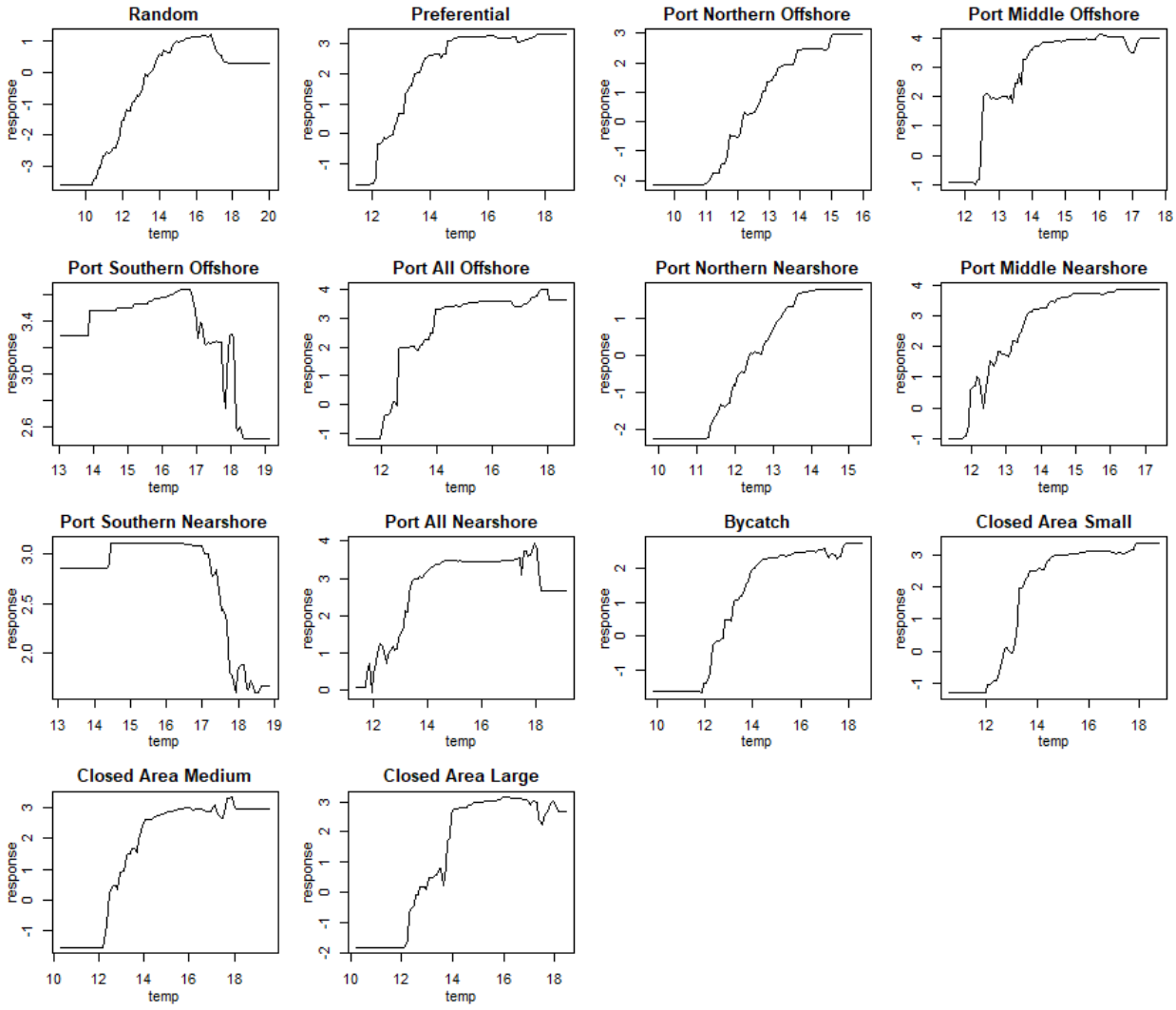

Figure S12. Sea surface temperature response curves for the binomial (occurrence, presence-absence) part of delta model (BRTs).

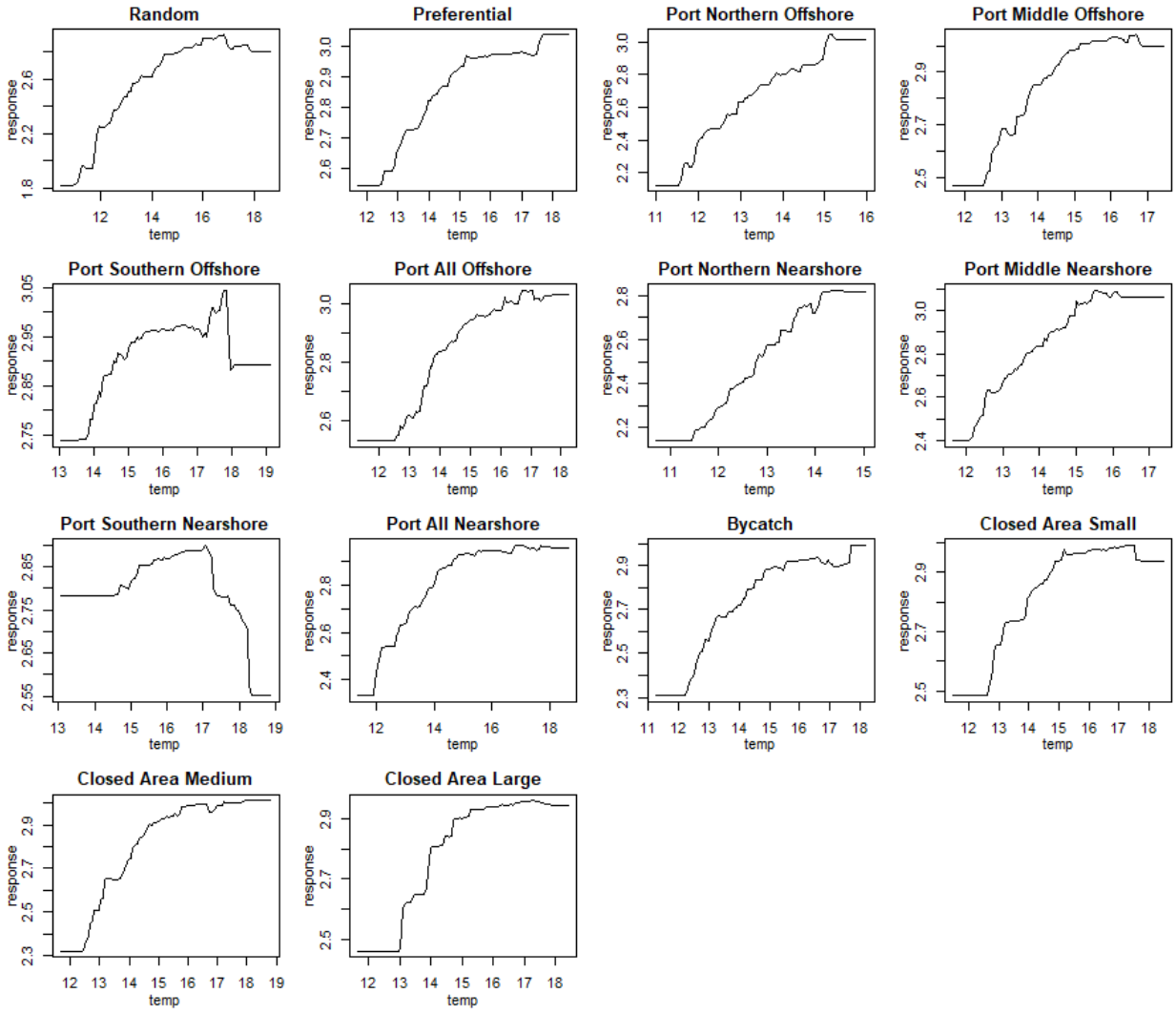

Figure S13. Sea surface temperature response curves for the abundance part of delta model (BRTs).

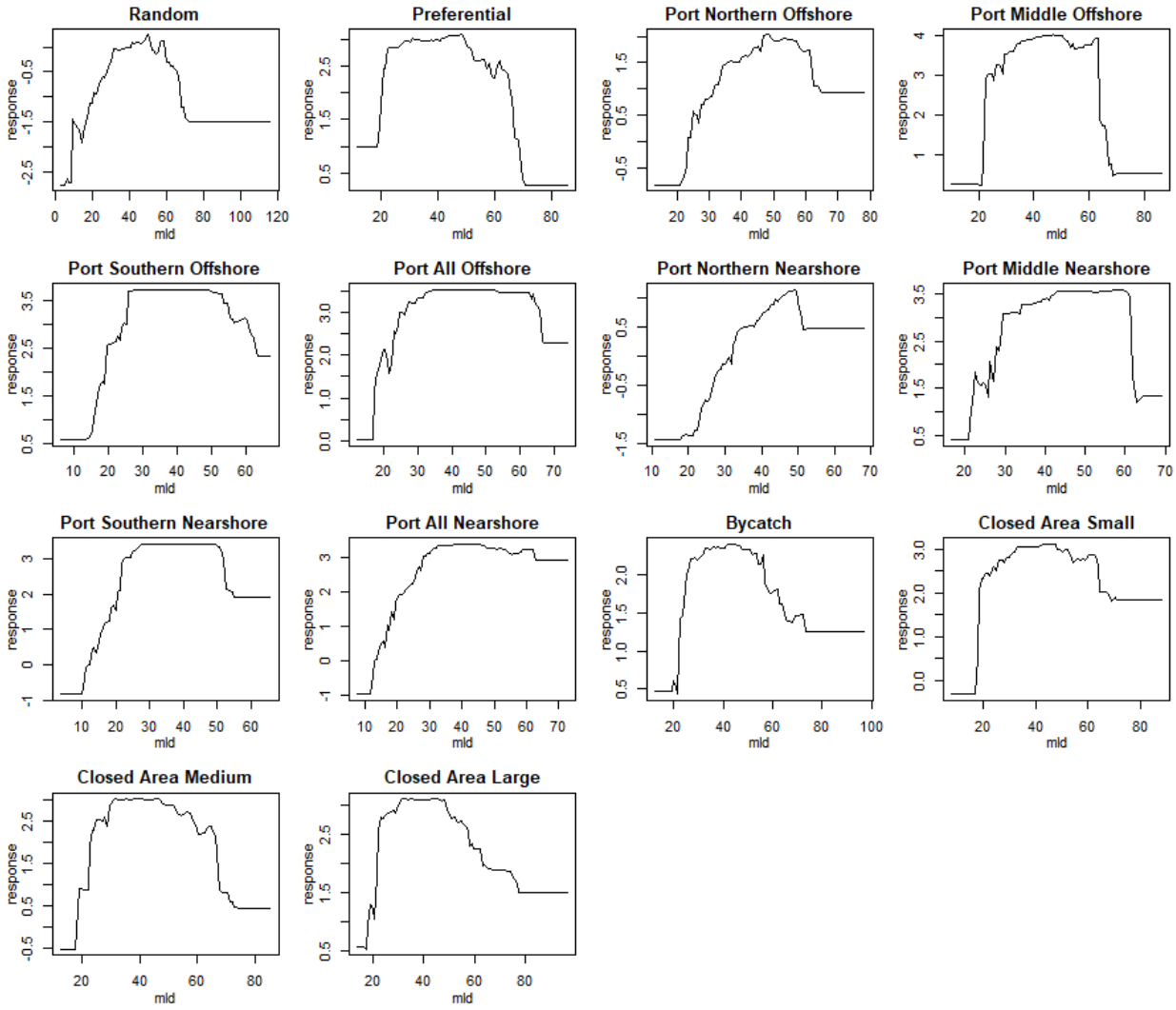

Figure S14. Mixed layer depth response curves for the binomial (occurrence, presence-absence) part of delta model (BRTs).

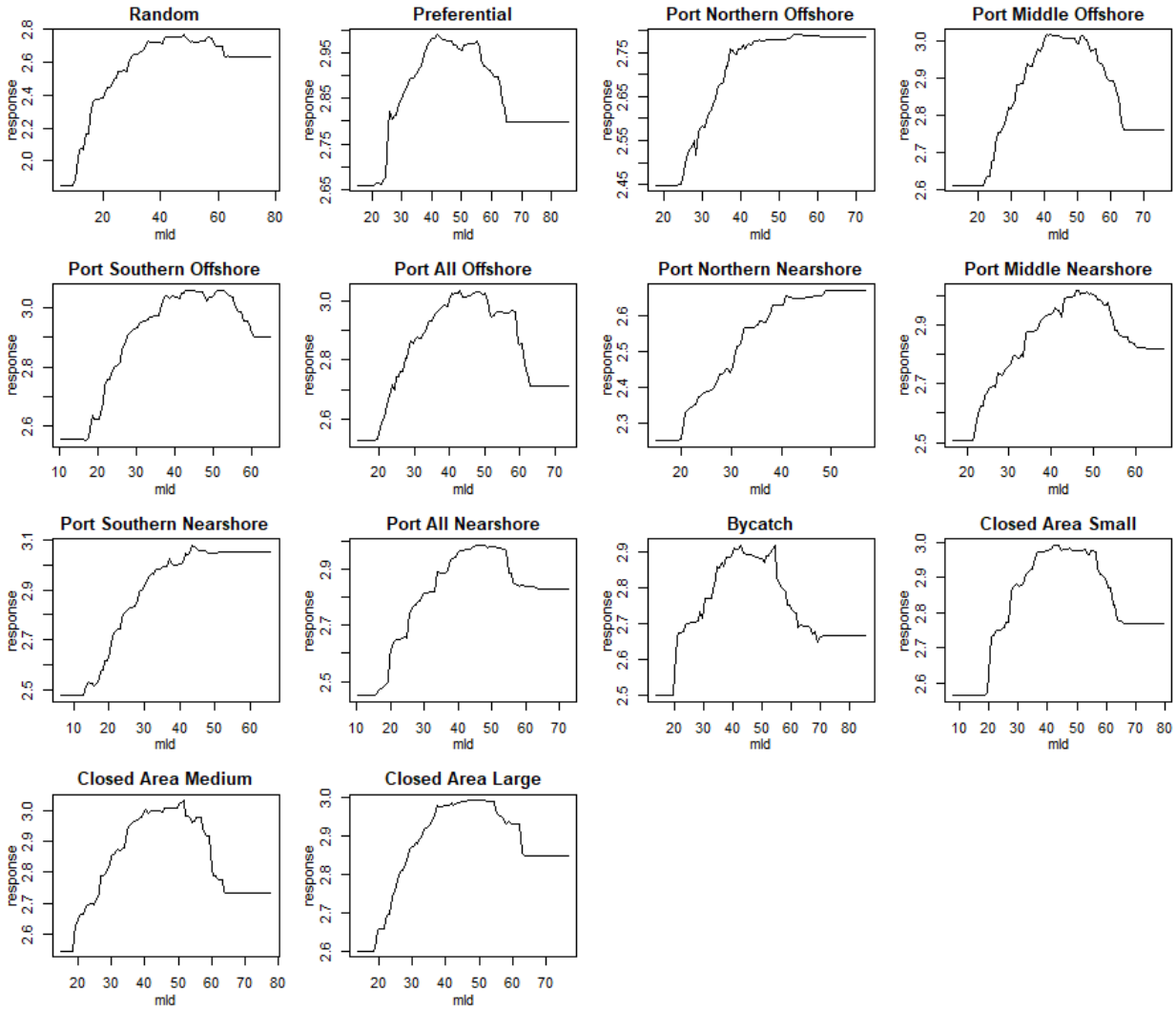

Figure S15. Mixed layer depth response curves for the abundance part of delta model (BRTs).

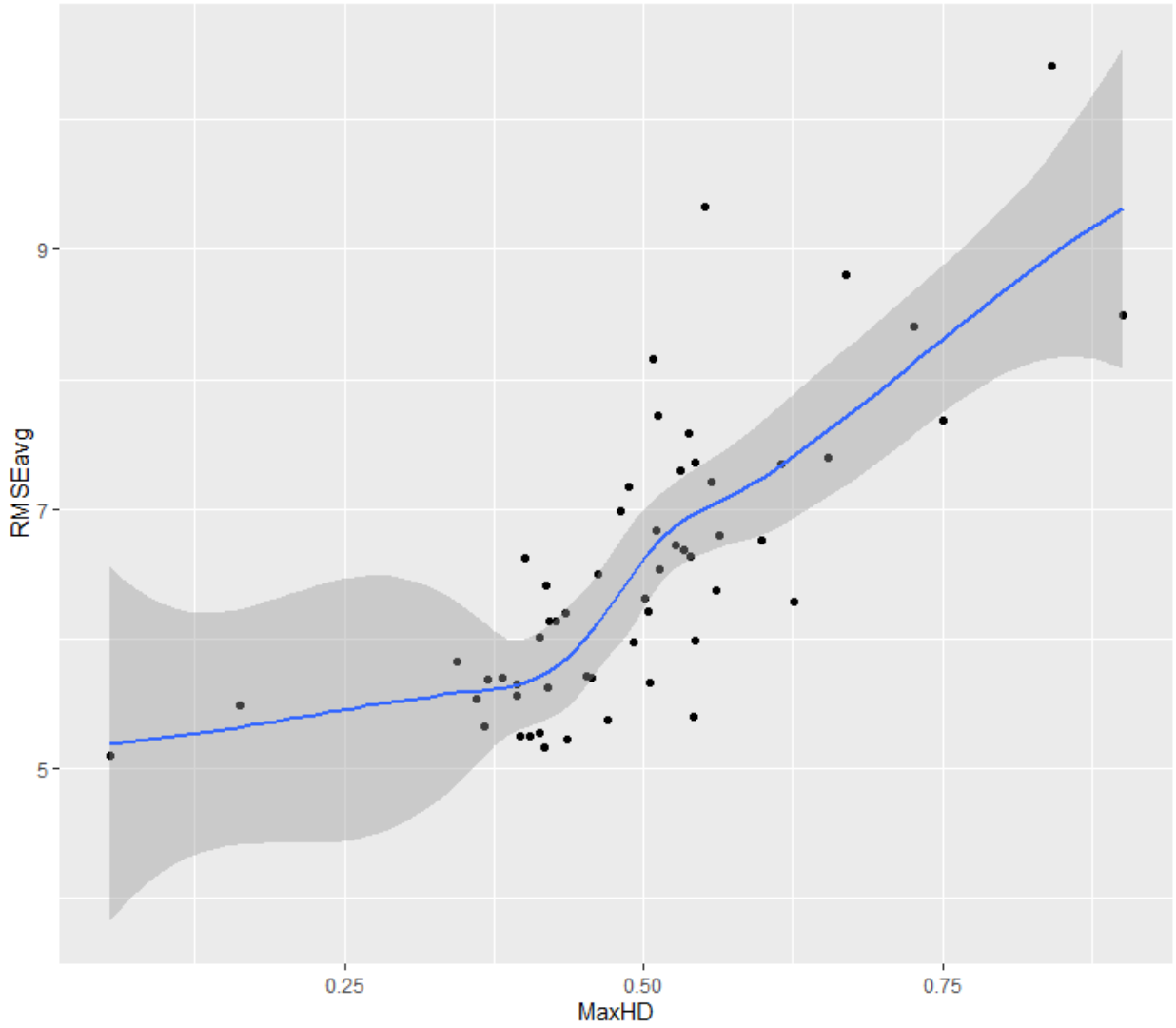

Fig S16: Comparison of Hellinger Distance (taken as the maximum HD across all climate variables for each sampling scenario and time period) and Mean RMSE.

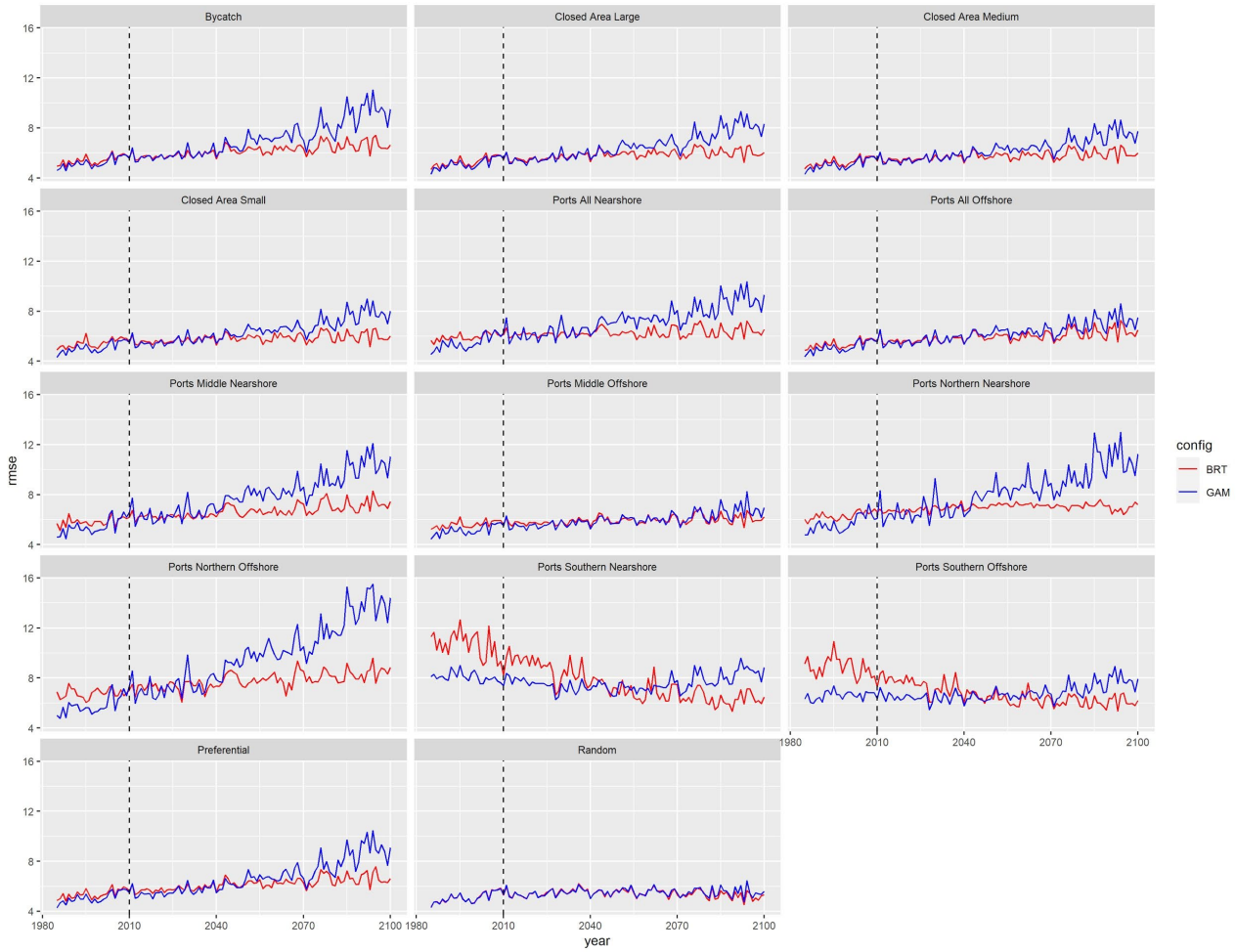

Figure S17: RMSE time series for models for 1985-2100. Each panel represents one of the 14 different fishing location scenarios, and each line represents the different model algorithm (GAM or BRT). Dashed line represents when the historical model fitting ends (1985-2010) and forecasts begin (2011-2100).

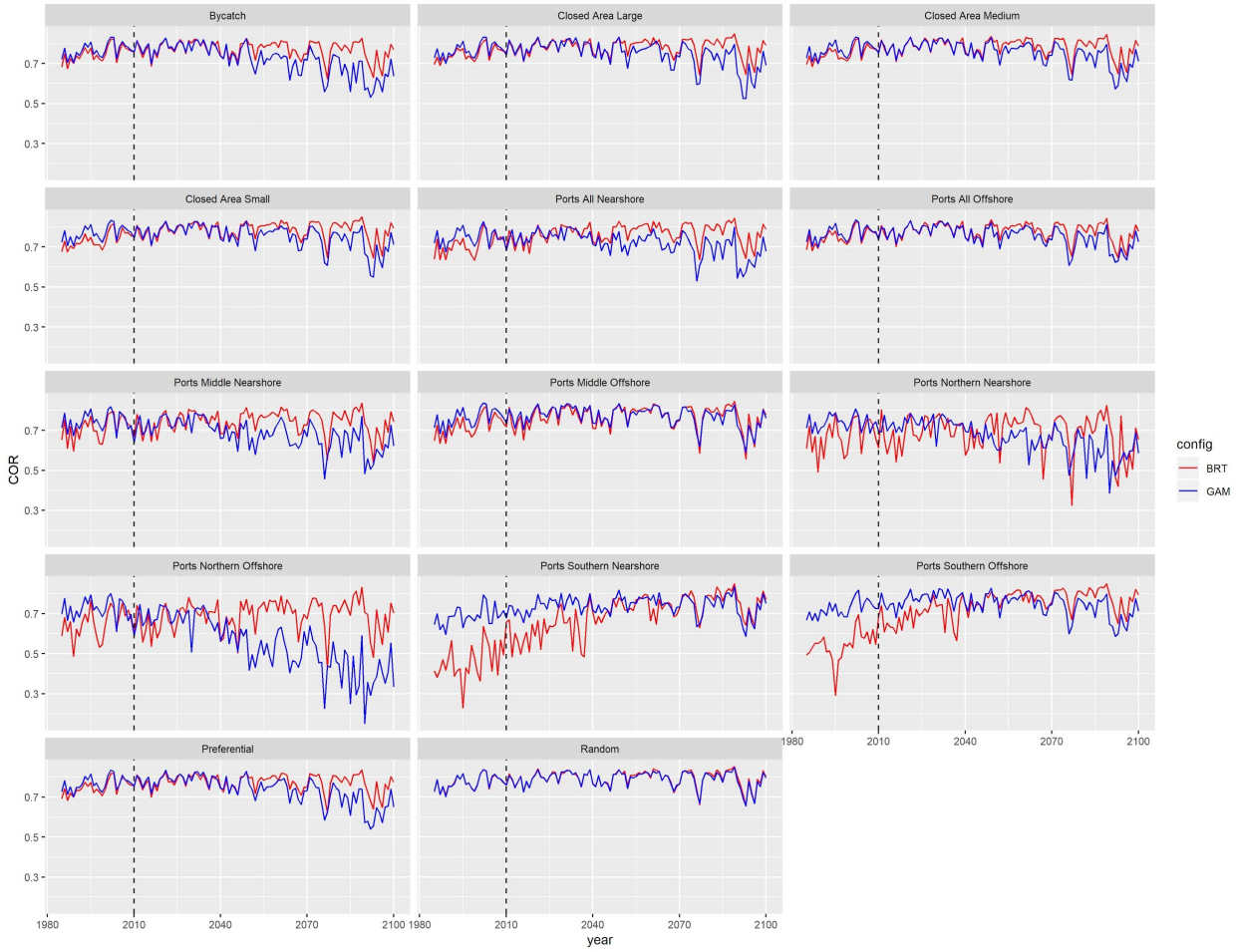

Figure S18: Time series of annually averaged correlation coefficients. Each panel represents one of the 14 different fishing location scenarios, and each line represents the different model algorithm (GAM or BRT). Dashed line represents when the historical model fitting ends (1985-2010) and forecasts begin (2011-2100).

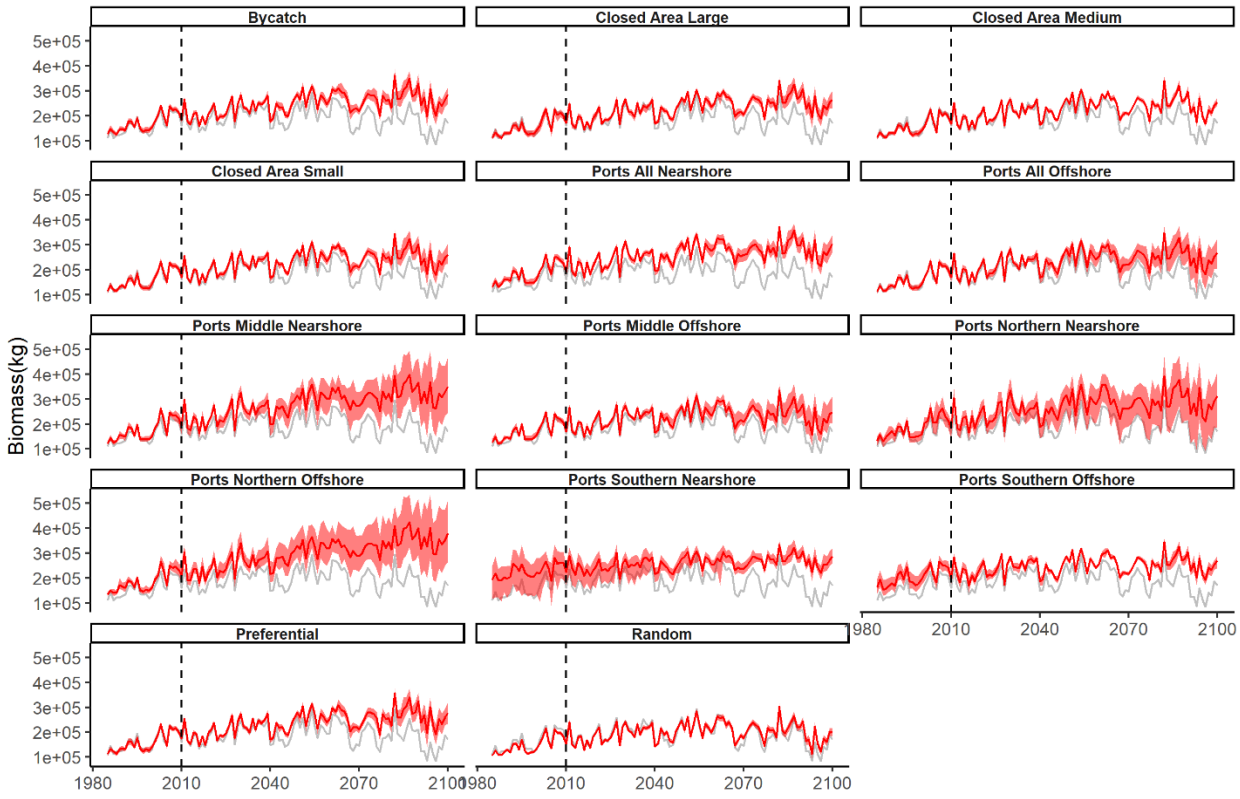

Figure S19: Time-series of simulated (grey line) and estimated (red line) biomass, with the within-model uncertainty for GAMs indicated by red shading. Results shown for each sampling scenario.

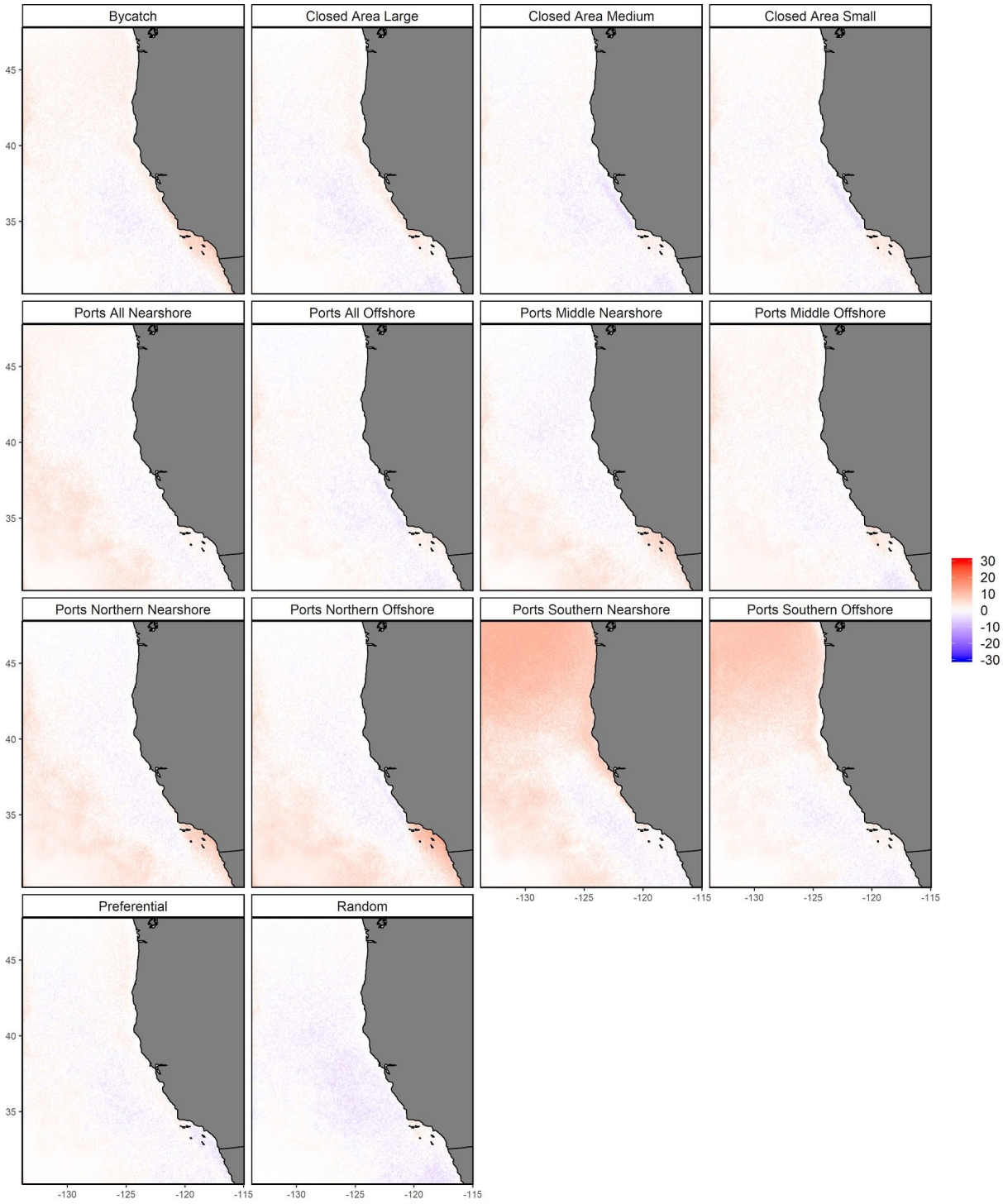

Figure S20: Difference between predicted and true abundance at each spatial grid cell for each sampling scenario for the historical (1985-2010) time period. Blue areas indicate areas where the model underpredicts the true abundance, and red areas represent the areas where the models overpredict the true abundance.

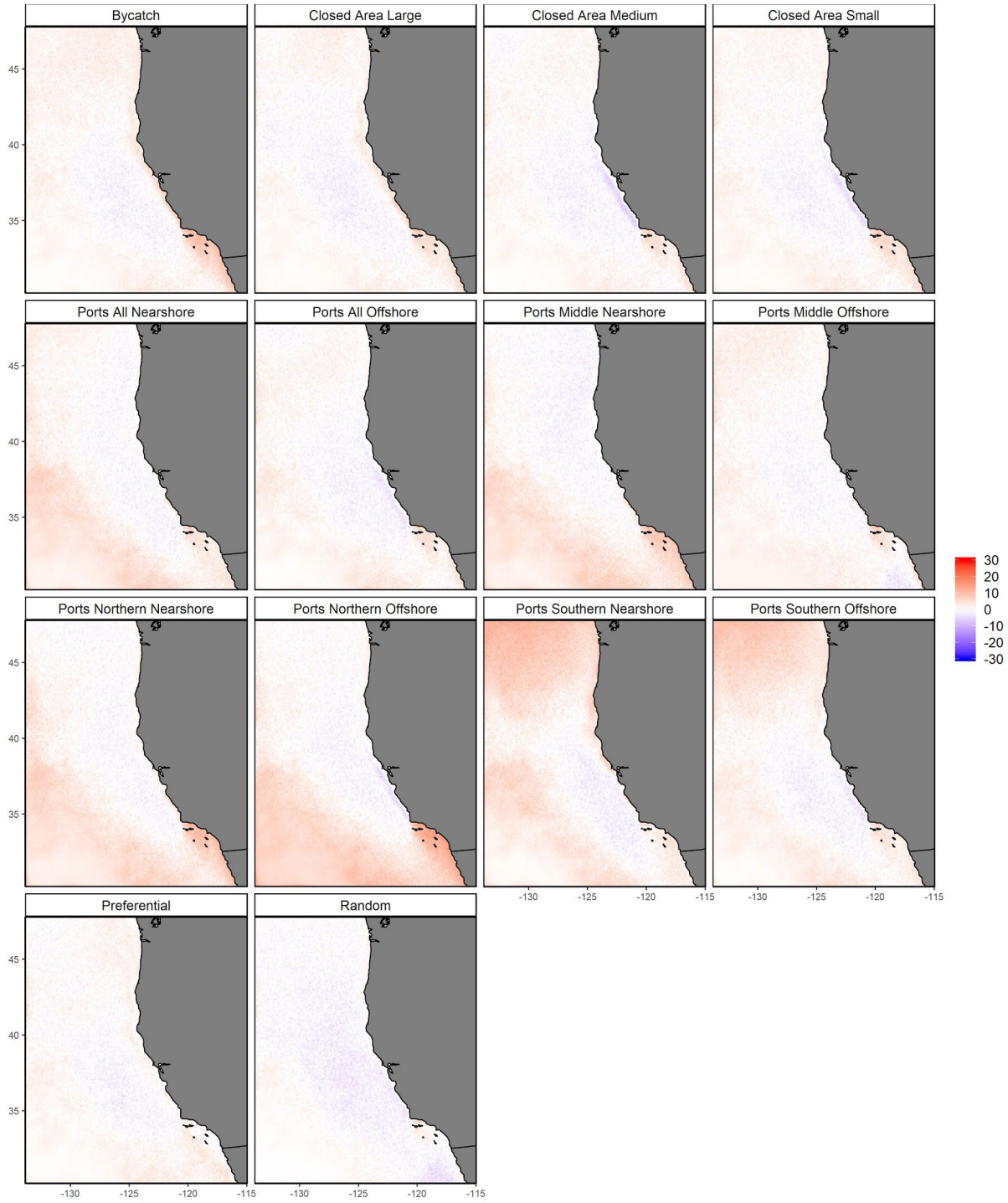

Figure S21: Difference between predicted and true abundance at each spatial grid cell for each sampling scenario for the early century (2011-2039) time period. Blue areas indicate areas where the model underpredicts the true abundance, and red areas represent the areas where the models overpredict the true abundance.

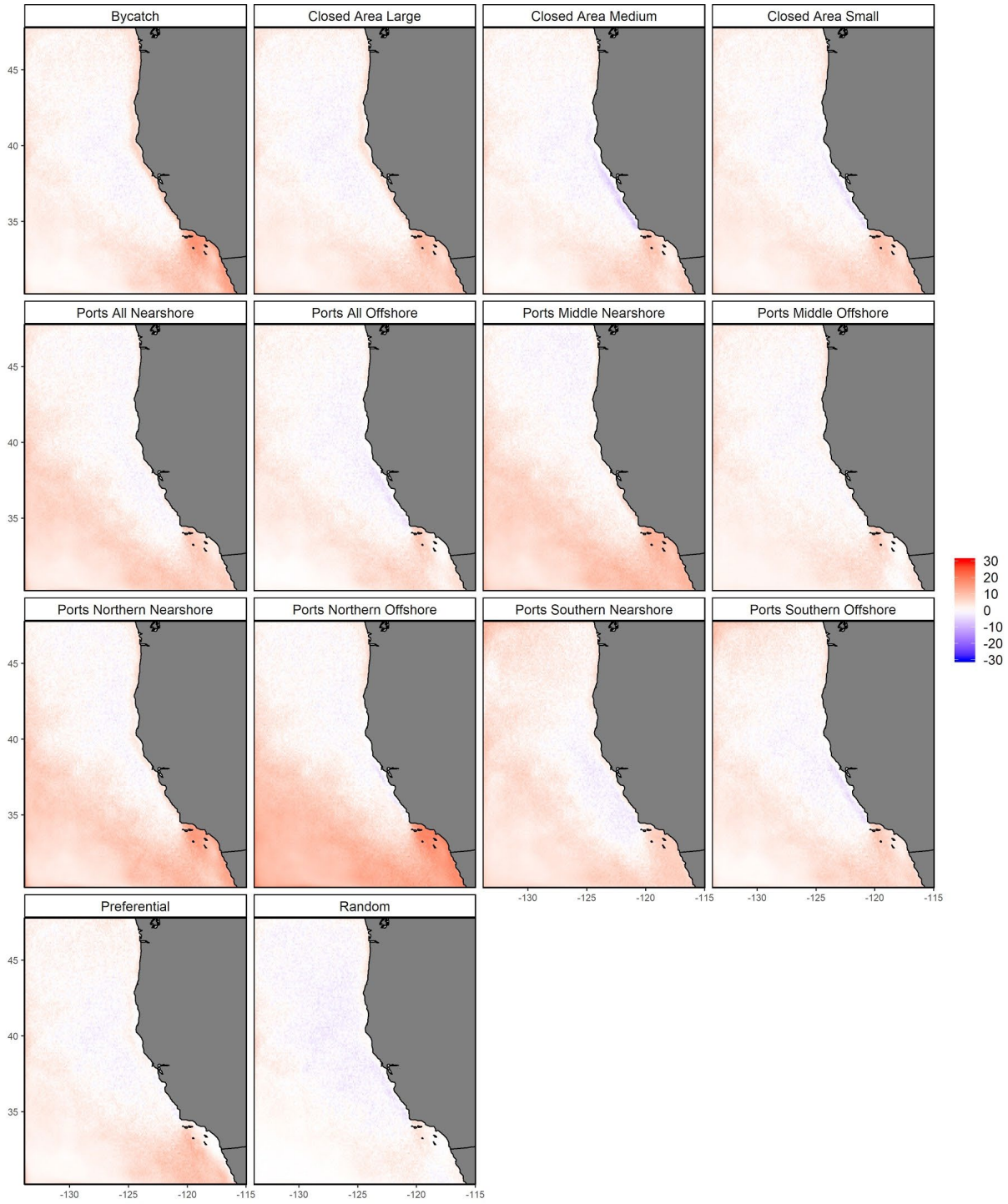

Figure S22: Difference between predicted and true abundance at each spatial grid cell for each sampling scenario for the mid-century (2040-2069) time period. Blue areas indicate areas where the model underpredicts the true abundance, and red areas represent the areas where the models overpredict the true abundance.

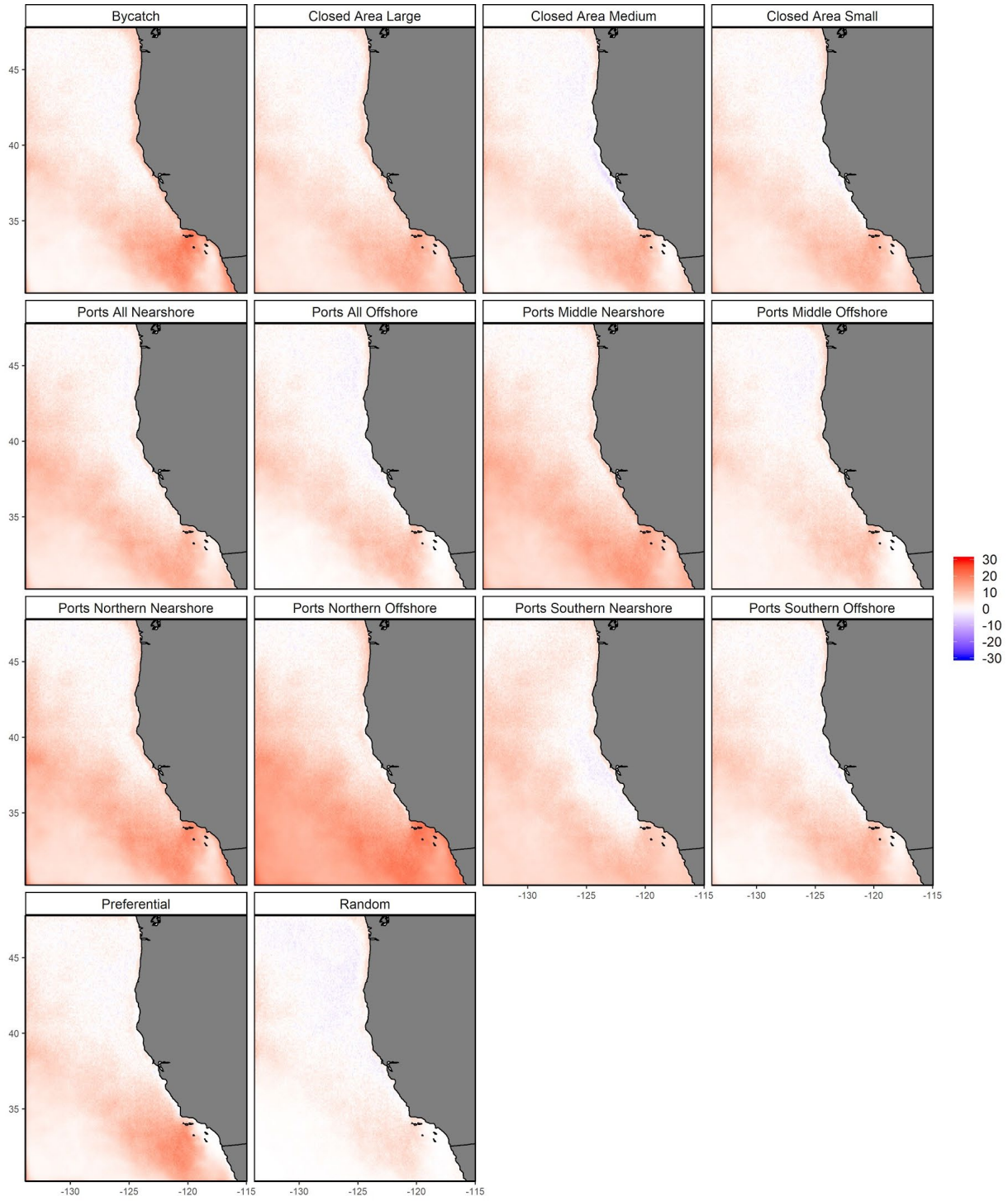

Figure S23: Difference between predicted and true abundance at each spatial grid cell for each sampling scenario for the late-century (2070-2100) time period. Blue areas indicate areas where the model underpredicts the true abundance, and red areas represent the areas where the models overpredict the true abundance.